DOCUMENT RESUME

ED 266 189 TM 860 135

AUTHOR Turnbull, William W.
TITLE Student Change, Program Change: Why SAT Scores Kept Falling.
INSTITUTION College Entrance Examination Board, New York, N.Y.; Educational Testing Service, Princeton, N.J.
REPORT NO CBR-85-2; ETS-RR-85-2
PUB DATE 85
NOTE 15p.
AVAILABLE FROM College Board Publications, Box 886, New York, NY 10101 ($5.00).
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Change Agents; *College Entrance Examinations; Educational Change; Educational Environment; *Educational Trends; Elective Courses; Enrollment Influences; Grades (Scholastic); High Schools; High School Seniors; Homework; Public Policy; *Scores; Social Influences; Standardized Tests; *Test Interpretation; *Test Norms; Test Results; Textbooks
IDENTIFIERS *Scholastic Aptitude Test; *Test Score Decline

ABSTRACT

This paper analyzes Scholastic Aptitude Test (SAT) score declines of two decades--the 1960s and 1970s. The importance of pervasive societal influences is accepted. A variety of data suggests, however, that the increase in school retention rates of poorly prepared students and the resulting heterogeneity of the high school population is a unifying exploratory variable for the second leg of the decline as well as the first. The argument is that several of the "pervasive influences" invoked to explain the continuation of the decline in the 1970s are best understood as adaptive responses of the schools to the appearance of a greater diversity of students in senior high schools. These responses, which in combination represent a reduction in the demand level of the school program, included grade inflation, proliferation of electives, textbook simplification, and reduction in homework assignments. Students also chose fewer academic and more vocational and general courses. It is hypothesized that the continuation of the decline in the 1970s was, in substantial part, the direct consequence of those school-related changes and thus was a delayed, indirect consequence of the compositional shift.
(Author/LMO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

College Board Report

No. 85-2

ED266189

Student Change, Program Change: Why SAT® Scores Kept Falling

William W. Turnbull

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. W. Gardner

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Student Change, Program Change: Why the SAT® Scores Kept Falling

William W. Turnbull

Educational Testing Service

College Board Report No. 85-2

ETS RR No. 85-28

College Entrance Examination Board, New York, 1985

William W. Turnbull is Distinguished Scholar in Residence, Educational Testing Service, Princeton, New Jersey.

Researchers are encouraged to express freely their professional judgment. Therefore, points of view or opinions stated in College Board Reports do not necessarily represent official College Board position or policy.

The College Board is a nonprofit membership organization that provides tests and other educational services for students, schools, and colleges. The membership is composed of more than 2,500 colleges, schools, school systems, and education associations. Representatives of the members serve on the Board of Trustees and advisory councils and committees that consider the programs of the College Board and participate in the determination of its policies and activities.

Additional copies of this report may be obtained from College Board Publications, Box 886, New York, New York 10101. The price is $5.

Copyright © 1985 by College Entrance Examination Board. All rights reserved.
College Board, Scholastic Aptitude Test, SAT, and the acorn logo are registered trademarks of the College Entrance Examination Board.
Printed in the United States of America.

CONTENTS

Abstract 1

Introduction 1

What the SAT Measures 1

Public Concern 1

Test Scores and the Schools 2

The Score Decline as Seen in the 1970s 2

Compositional Change 3

Pervasive Change 4

Problems with Unrelated Declines 4

The "Reciprocal Score Declines" Alternative 5

Societal Factors 5

Textbooks 5

Electives 6

Choice of Curriculum 6

Homework 6

Grade Inflation 7

Fewer High Scores 7

Scores of High-Ranking and Low-Ranking Students 7

Earlier Grades 8

Compositional Change Had a Pedagogical Echo 8

What Can the Decline Teach Us? 9

The Decline Is Probably Over 9

Concern about Overemphasis on Academic Proficiency 9

Needed: A System to Serve Diversity 10

References 10

Tables

1. SAT Verbal and Mathematical Score Means, 1962-63 and 1976-77 2

2. Mean SAT Scores by Class Rank (1976, 1980, 1984) 8

ABSTRACT

The first leg of the Scholastic Aptitude Test (SAT) score decline occurred mainly in the 1960s. It seemed to be explained fairly satisfactorily by the evidence that the composition of the test-taking group had changed to include a larger proportion of students with relatively low-developed ability, mirroring the increased holding power of education for teenagers. In studies made during the 1970s, no comparable underlying change was found to explain the second (mainly 1970s) segment of the decline, which was ascribed instead to a complex of factors—"pervasive influences"—in both school and society.

In this study, the importance of pervasive societal influences on student learning is not in dispute. A variety of data suggests, however, that the increase in school retention rates of poorly prepared students and the resulting heterogeneity of the senior high school population is a unifying explanatory variable for the second leg of the decline as well as the first. The argument is that several of the "pervasive influences" invoked to explain the continuation of the decline in the 1970s are best understood as adaptive responses of the schools to the appearance of a greater diversity of students in senior high school. These responses, which in combination represent a reduction in the demand level of the school program, included grade inflation, proliferation of electives, textbook simplification, and reduction in homework assignments. Students also chose fewer academic and more vocational and general courses. It is hypothesized that the continuation of the decline in the 1970s was, in substantial part, the direct consequence of those school-related changes and thus was a delayed, indirect consequence of the compositional shift.

The two declines seem to have been reciprocal rather than unrelated. Student changes begat program changes, and each new condition in its turn led to lower scores. The paradigm helps to explain why high scores became scarcer in the 1970s, why the scores of students with high class rank continued to drop while those of students with lower class standing stabilized or turned up, and why the pervasive factors seemed to affect only the senior high school students.

This report suggests that the academic demand level of the schools probably stopped falling in the late 1970s, that 1980 marked a significant turning point, and that the recent slight upturn in SAT scores marks the beginning of a positive trend rather than a pause before a continuation of the drop.

INTRODUCTION

In the 1970s the nation discovered a new barometer to use in gauging America's educational climate: the SAT score averages reported yearly by the College Board. The news was bad. The scores had gone down steadily since 1963 and, in the early 1970s when this trend was first noted, the drop was becoming steeper rather than leveling off. Widespread public concern developed quickly and the score decline became the subject of long and intense debate, which continues today.

Many observers seemed ready to take the SAT score decline as an index of the quality of schooling, over the strong objection of others who pointed out the inappropriateness of doing so. Those who reject the SAT as a barometer of schooling are on firm ground. Students who take the test are a representative sample of neither high school seniors nor college-bound students. The SAT was never intended to represent all of the important areas of understanding, knowledge, or skill—not to mention constructive attitudes, values, and other noncognitive characteristics—in which schools aim to bring about student growth. Moreover, the scores are not affected only by formal schooling: they measure abilities that are developed both in and out of school.[1]

Although on reflection most people are ready to accept the limitations of SAT scores as a sufficient indicator of the quality of schooling, few are willing to dismiss the score trends as irrelevant to education.

WHAT THE SAT MEASURES

The SAT measures a student's proficiency in solving word and number problems. The intellectual tasks posed by the test—difficult but not impossible for most students who are applying for college admission—require the application of skills that are useful in academic work and that are developed at least in part through formal education. The test stresses higher-order reasoning rather than the "basics" of reading and calculating. On the average, students who receive high scores do better academically in both school and college than those who score less well. The test is taken by about a third of all high school seniors and by over two-thirds of all college entrants—a large group of national significance. Colleges have validated the correlation between success on the SAT and success in college, and they obviously feel that it is important for high school students who want to go on to college to have developed the skills represented by good SAT scores.

PUBLIC CONCERN

For all of the reasons just listed, parents and taxpayers care very much about whether the SAT averages are rising, falling, or holding steady, and they want the schools to help keep the scores up.

1. The College Board has spoken strongly and consistently against attempts to use SAT scores to measure American education. See the *College Board News* (Spring 1984) for a summary of Congressional testimony on this point by Daniel Taylor, Senior Vice President.

Public dissatisfaction over the long and wearying decline of test scores in the 1960s and 1970s was a powerful stimulus for change. In the late 1970s it was a mobilizing force behind the enactment of a wide variety of state and local movements to reform education. As a secondary consequence, it paved the way for the appearance in the early 1980s of nationally prominent reports on "the crisis in the schools." These reports largely followed rather than led the grass-roots concern but they served to validate the slippage in student attainment and helped to set the rhetoric for discussing it, as well as offering prescriptions for constructive change.

From 1963 to 1980 the mean SAT scores went down by 54 points on the verbal and 36 points on the mathematical sections. Since 1980 they have held steady and even edged up a little: two points on the verbal and five points on the mathematical sections from 1979–80 through 1983–84. So closely are they now being watched that this small rise was hailed at the national level in September 1984 as signaling the turnaround of American education! In Washington, Secretary of Education Terrel H. Bell attributed the increase to "the movement toward excellence in our schools that is sweeping the nation" (*Chronicle of Higher Education* 1984, p. 11). That may be true. It seems likely, however, that whatever change has occurred since 1980 owes more to parental concern and local school action in the 1970s than to the national studies, however useful, that have appeared in the last two years.

TEST SCORES AND THE SCHOOLS

It is by no means the intention of this report to side with those who would accept the SAT as a proper barometer for American schooling. Such an equation is grossly misleading. It would be unfortunate, however, if we were led to the opposite conclusion: that because the scores are not a sufficient index of what is happening to all the skills of all the students, they can tell us nothing worth knowing about some of the skills of some of the students.

The view that the score decline is important and should be examined further is reinforced by the fact that the SAT score trend is consistent with similar trends in other data. Scores on the ACT program tests at the senior level have shown a similar pattern, as have scores on the National Assessment of Educational Progress given to 17-year-olds (in contrast to the smaller declines shown at the 13-year-old level and, in some cases, gains by the 9-year-olds). Recent corroborating evidence comes from a comparison of data from two representative samples of high school seniors studied in 1972 (National Longitudinal Study) and 1980 (High School and Beyond) under the direction of the National Center for Education Statistics. Fetters, Brown, and Owings (1984, p. 20) note that "the declines of one-fifth standard deviation in verbal and one-sixth standard deviation in mathematics observed in this study are about the same sizes as the declines in SAT scores (one-quarter and one-seventh standard deviation, respectively) from 1972 to 1980." Although the SAT scores are in no way representative of the entire high school class, their pattern of change does not conflict with the data reported from other sources.

It is axiomatic that SAT scores or any other measures of what young people learn while growing up are strongly influenced by what is happening in their homes and in the society in which the students and their schools are embedded. And yet it is plausible that, as the public and many commentators obviously believe, the trends in scores do bear a relation to changes in the formal education system. A decline as sweeping as the one we have seen in a generation presents us as educators with an obligation to explore the educational lessons that we may be able to learn from it. At a minimum, such an examination may help us understand the past; at best, it may provide some guidance as we weigh the variety of reforms that are proposed for the future.

THE SCORE DECLINE AS SEEN IN THE 1970s

When the persistent drop in SAT scores was first announced by the College Board in 1975, the public alarm coupled with the magnitude of the change was strong enough and worrisome enough to lead the College Board to invite an Advisory Panel on the Scholastic Aptitude Test Score Decline to look into its causes.

In 1977, the Panel, chaired by Willard Wirtz, issued a thoughtful report, entitled *On Further Examination* (College Board 1977), which examined the decline and its possible causes in detail. The statistics that provided the broad outline of the decline are presented in Table 1. These numbers meant that on the verbal sections of the test, where scores had dropped 49 points, only about one-third of the SAT takers in 1977 could answer as many questions correctly as the average test taker in 1963 could answer. On the mathematical sections the decline of 32 points was not quite so severe as the drop in verbal scores but it still was steep enough for concern. In 1977, around 40 percent of the students could match the 1963 mathematical average.

Examination of other data reported by the College Board shows, moreover, that the average *senior* taking the

Table 1. SAT Verbal and Mathematical Mean Scores, 1962–63 and 1976–77

	SAT-V	*SAT-M*
1962–63	478	502
1976–77	429	470
Decline	49	32

SAT in 1977 did no better on the verbal sections (with a mean score of 429) than the average high school *junior* had done in 1963–64, as measured by the Preliminary Scholastic Aptitude Test/National Merit Scholarship Qualifying Test (PSAT/NMSQT) given in that year.[2] The mean of the 1977 senior mathematical scores (470) stood only five points above the mean PSAT/NMSQT performance of the 1963 juniors in mathematics (465). On both the verbal and mathematical sections of the test, then, the scores of high school seniors had slipped to a degree that most people found alarming. Granted, the test takers were not representative of all high school students, but they made up two-thirds of the college-bound population. The decline was something worth worrying about.

On Further Examination explored the evidence for and against a wide variety of hypotheses advanced to explain the decline: (1) changes in the nature of the SAT, in its relevance or in its score scale; (2) changes in the composition of the group taking the test; and (3) changes in the schools and in "the broader learning context"—the environment in which young people were growing up in the 1960s and 1970s.

The authors found little evidence to suggest that the score declines reflected the factors in category (1), that is, it did not seem to stem from changes in the score scale, in the test, or in the fit between the test and what was taught in school. With respect to the other possible causes—(2) composition of the group or (3) the learning environment—the Panel (College Board 1977) reached a bifurcated conclusion:

> Two Score Declines
>
> Fourteen years of uninterrupted decline in the SAT scores create the illusion that there is some single force or closely related set of forces at work here. This isn't the case. The decline has developed in two distinct stages, characterized by significantly different balances of materially different causal factors.
>
> During the first six or seven years of the decline the composition of the SAT-taking population was changing markedly. Each year it included larger proportions of characteristically lower-scoring groups of students. This pulled the overall average down. There were only slight falloffs during that period in the score means *within* any particular ability groups.
>
> The pattern changed after about 1970. The "compositional" shifts slowed down materially. What showed up increasingly was an across-the-board score decline, the apparent consequence of more "pervasive" changes or influences affecting higher- and lower-scoring groups alike (p. 13).

2. The PSAT/NMSQT data are taken from Jackson (1976b, pp. 2-3). The PSAT verbal mean score was actually 42.9 on its own scale but can be converted by multiplying by 10 to yield the equivalent SAT score of 429. A comparable adjustment is made in the case of the math score comparison immediately following.

COMPOSITIONAL CHANGE

The end of World War II had set in motion a series of remarkable transformations in American life and in American beliefs. Among the latter was an expanded sense of the centrality of education as preparation for postwar life—not for the privileged few but, as was commonly said, for "all the children of all the people." Particular attention was paid to the need for encouragement of children from poor and minority families to continue their education and prepare for better jobs as part of the effort to combat social and racial injustice and break the generational cycle of poverty. The high school diploma was seen not only as a proper goal but as an urgent need for all young people. This new priority, applied to the tidal wave of young people who made up the postwar baby boom, led to an unprecedented leap in the number of students who stayed to finish high school.

These facts were fully recognized by the Panel on the Score Decline in its report (College Board 1977):

> It is already hard to remember the extraordinary confluence of forces that struck the educational system in the 1960s.
>
> About 1,864,000 students graduated from high school in 1960, and some 564,000 juniors and seniors took the SAT that year. Ten years later the number of graduates had increased by a million, and the number of SAT takers had tripled.
>
> This was partly a demographic change. Between 1964 and 1965 the number of 18-year-olds in the country jumped by more than a million. That was when the post-World-War-II population wave first hit this age level. It was also the time the SAT average scores started down.
>
> By what was probably more than coincidence, the nation decided during that same period to reduce the high school dropout rate and to see to it that a larger percentage of young people had the opportunity to go on to college.
>
> Perhaps it was a historical accident that this was also the time of tardy legislative decision to attack previous discrimination in providing educational opportunity, particularly discrimination based on race, sex, and family income.
>
> Twenty-five years ago, only half of all young Americans were staying in school through the 12th grade; this fraction grew by 1964 to two-thirds and by 1970 to three-fourths. The proportion going on to college was about one-fourth in 1952, about a third in 1964, and almost half by 1970 (p. 14).

The "new students," as K. Patricia Cross has called them, were by no means duplicates of the traditional high school graduates and college applicants. In the main, their academic preparation was not as strong. Again, the Panel (College Board 1977) makes the point well:

> It would be pleasant to think that as increased percentages of vastly larger numbers of young people stay in school longer and go on to college, the college entrance examination averages achieved before by a favored fraction of students could be held constant. Yet any such expectation would be ruefully unrealistic. The major move toward

> equality of opportunity in the 1960s will be judged unfairly unless it is recognized that an increasing school retention rate is bound to mean, at least at first, some drop in the *average* developed ability level (as reflected in traditional tests) of the larger number staying the course.... Starting in about the mid-1960s, cumulatively larger *percentages* of students with comparatively lower high school grade averages were going on to college (p. 14).

This inevitable effect of the "compositional change" on SAT scores had in fact been discussed in the 1960s. In an article in *Science*, Turnbull (1968) pointed out:

> ...In the years from the end of World War I to 1960 we succeeded in expanding markedly the proportion of students in the top-quarter ability group who go on to college, to the point where we had very nearly reached the limit for that group. This means that all of the further increase in the proportion of students going beyond high school must come from the second, third and fourth ability groups:...these students are largely from the score range in the 200's and 300's.

As the Panel on the Score Decline noted, however, by the early 1970s the proportion of students going beyond high school had reached a plateau and the "compositional" shift in the SAT-taking population had slowed materially. The single most evident cause of the score decline seemed to have largely run its course. And yet the scores accelerated their downward path. The Panel saw not one single reason for this continuation but rather a variety of probable causes, several of them interwoven in a pattern of "pervasive change"—the set of circumstances that they invoked to explain why the scores did not stabilize when the test-taking population did.

PERVASIVE CHANGE

The elements of the pervasive change perceived by the Panel as implicated in the second stage of the decline were diverse. Some of them were school-related, for example,

- fewer basic courses and more electives;
- less emphasis on writing/more on objective tests;
- grade inflation and lower standards;
- automatic promotion from one grade to the next;
- teachers who were less well-prepared.

Some were largely beyond the reach of the school, for example,

- more time spent watching television;
- decline in family influence: parents not effective as teachers;
- the war in Vietnam, and draft policy;
- a decade of distraction.

And for some the responsibility was probably mixed, for example,

- greater absenteeism;
- less time spent on homework;
- a decline in student motivation to learn.

The grouping of these possible ingredients of the problem is mine rather than of the Panel, which went on in its report to mention a dozen other specifics, from birth order to nuclear fallout, that were being put forward as possible causes for falling test scores.

This list, read seven years after the Panel's report, is likely to elicit nods of agreement from those who lived through the period. It seems extremely likely that most or all of the several influences believed by the Panel to account for the score drop in the 1970s were indeed contributors to the decline. Since a number of these factors—for example, the war in Vietnam—lay beyond the school, some people have concluded that the diminishing achievement of students in the 1970s was a negative report card on the nation, not on the schools, especially since the SAT scores are neither specific nor sufficient measures of school effectiveness. But there are problems in seeing the decline as reflecting two essentially unrelated sets of phenomena—the first compositional, the second pervasive in school and society--that happened to occur back to back.

PROBLEMS WITH UNRELATED DECLINES

The "compositional shift" explanation for the first stage of the decline was widely accepted. It was intuitively satisfying, it was consistent with the evidence, and it allowed us to congratulate ourselves on our dramatic progress toward equality in American secondary education. The explanation of the second stage as resulting from "pervasive changes," however, somehow seemed less satisfactory to many people. They were, it seemed, looking for a cause for the second decline as encompassing and perhaps as comforting as the "compositional shift" that provided a plausible rationale for the first decline.

There was also the fact that the pattern traced by the SAT mean scores from 1963 through 1976 showed no sign of a letup or break in the middle. The scores had continued to fall even more rapidly after the compositional shift had stopped. Clearly, then, the new factors, if that was the explanation, must have come along at just the right time to take up where the compositional shift left off.

The timely emergence of the suggested causes of the "second leg" of the decline may have been coincidental—the causes may have been unrelated—but it seems more likely that there were important connections between them that help to explain the continuity evident in the score trends.

THE "RECIPROCAL SCORE DECLINES" ALTERNATIVE

In brief, the compositional shift itself may have helped set in motion the set of circumstances that led to the SAT score decline over the full period from 1963 to 1980. The "reciprocal score declines" hypothesis which I propose accepts the proposition that the score changes up to the early 1970s were an early and direct reflection of the fact that the schools were retaining into grade 12 more students whose achievement was marginal by previous standards, that more of these students were applying to college, and that a large proportion of the colleges were requiring the SAT for admission. The "reciprocal declines" theory goes on to suggest, however, that the score changes since that time have been, to an important degree, a later and indirect reflection of the same phenomenon: the new composition of the upper high school grades.

The new student body as it evolved in the 1960s led to modified pedagogic strategies, geared to the fact that there were now more students who found the traditional program too difficult. These strategies took cumulatively greater effect through the 1970s and played a large part in bringing about the continued score decline at both high and low score levels. In this sense, the compositional shift was the single major cause of the full extent of the decline that began in the 1960s, that continued to about 1980, and that now seems to have bottomed out and perhaps reversed. We had started a cycle in which a more heterogeneous student body, with more students who were poorly prepared, led to a lower level of expectation and demand in school, which served to feed the decline in performance on the part of the class as a whole, including the increasingly bored and understimulated students of medium and high ability. The decline in performance led to a still further relaxation of requirements, and so on in a continuing downward spiral.

The change in the student body led to a change in the school's program, and each of these changes, in its time, led to lower scores.

SOCIETAL FACTORS

The downward spiral obviously could not and did not occur as a phenomenon isolated within the schools. The Score Decline Panel was right in seeing the decline as consistent with and reinforced by the priorities of the larger society. The fact that SAT-verbal scores dropped more precipitously than SAT-mathematical scores is consistent with this view. Verbal scores are clearly a joint reflection of both in-school studies and out-of-school pursuits. Although mathematical scores are certainly not impervious to what happens out of school, they are more directly influenced by specific instruction. The drop in mathematical scores, then, is likely to be a purer reflection of the school-related changes than is the drop in verbal scores—and the mathematical score decline was just two-thirds as large as the verbal score decline.

The direction of the mathematical versus verbal discrepancy and its size should not be overinterpreted in relation to the comparative effects of school and society on the SAT scores, but the fact that the mathematical decline, largely reflecting school, was notably smaller than the verbal decline, reflecting both the school and the culture, helps to confirm the thesis that the in-school and out-of-school changes were working in the same direction. I believe, however, that we would miss the educational significance of a startling change in student performance—the dramatic decline in both scores—if we failed to look with care at the evidence for what was happening *in the schools* as well as what was afoot beyond their walls.

In the next sections of this report I shall cite a number of observations relevant to the hypothesis that the second leg of the connected declines reflected changes made by the schools to adapt to changes in the students, especially in grade 12.

TEXTBOOKS

One critical change directly related to schooling is the simplification of high school textbooks that took place in the 1960s and early 1970s. This change was a predictable response to the fact that a larger group of high school students found the traditional texts both difficult and boring. Jeanne Chall's (1976) study found that, "The history, literature, and grammar and composition eleventh-grade textbooks had an overall readability level of 9-10 grade level," and asked, "Can pupils who are not exposed in their classrooms during their eleven years of schooling to reading matter as difficult as the SAT reading selections be expected to do well on them?"

The Chall study of 1977 marked the beginning rather than the end of the textbook debate, which continues today. "In recent weeks," reported Fred Hechinger in a column in the *New York Times* on April 10, 1984, "the demand for better textbooks has joined the chorus of school reform proposals. Last month, high-level educators, called together by Governor Robert Graham of Florida, responded to what Secretary of Education T. H. Bell had earlier called 'the dumbing down' of textbooks. . . . The decline of textbooks is a chicken-egg problem. Schools teach or demand less, and pretty soon pupils need easier books." The cycle is thus perpetuated.

Most of the criticism of textbooks has been directed toward those used in literature and history classes. No comparable concern is voiced about changes in the content of high school mathematics texts. The readability formulas applied to such texts by school districts considering a book purchase (and therefore applied in advance by the publishers) relate especially to vocabulary and the length and complexity of sentences. One may speculate that the level of

demand has declined more generally in the verbal than in the mathematical textual material used in high school and that the difference has contributed to the steeper decline in SAT-verbal than in SAT-mathematical scores.

ELECTIVES

Making the textbooks easier was only one answer to the needs of less proficient students. Another was the addition to the curriculum of electives through which students could bypass some of the core courses if they found them too difficult.

The Score Decline Panel (College Board 1977, p. 28) noted a fact that may relate to the steeper drop in SAT-verbal than in SAT-mathematical scores. "It is perhaps relevant that there has not been the proliferation of electives in high school mathematics that has been characteristic of the verbal skills area. While Massachusetts high schools were increasing their course offerings in English by 50 percent (between 1971 and 1976) there was virtually no increase in the number of Mathematics courses being offered."

It is hard to say how much the proliferation of electives in the 1960s and 1970s had to do with the SAT score decline. We do not have good data on the extent to which the SAT-taking population shifted toward easier electives: the shift was probably greater within the segment of the senior class that did not take the SAT. It is likely, however, that some students at every level of academic performance were enrolled in relatively undemanding elective courses.

For the college-bound student there was a countertrend toward one kind of harder "elective": the Advanced Placement Examinations of the College Board. Advanced Placement (AP) enrollments gained rapidly from 1963 (29,000 examinations taken) to 1969 (69,000), plateaued through 1973 (71,000), and resumed their dramatic rise through 1980 (160,000) to 1984 (240,000). This remarkable increase reflects in part the rise in the number of schools offering Advanced Placement courses (from about 1,700 in 1963 to 6,300 in 1984). In 1984, 177,000 students took the 240,000 AP examinations given—less than a tenth of all seniors but closer to a fifth of those going directly to college. This is an encouraging development. By offering Advanced Placement classes, schools were providing able students with a positive form of curricular enrichment to counter the downtrend in the challenge presented in simplified core courses or in easier electives. The rising student response indicates the readiness of many students to move at a faster academic pace. It is still true, however, that the 6,300 schools offering AP courses comprise a distinct minority of high schools offering grade 12 work. Relatively few small schools, in fact, offer AP courses.

It is noteworthy that in 1980 a higher percentage of high school seniors were in remedial English courses (31 percent) than in advanced courses (27 percent). The same was true in mathematics, with 30 percent of seniors in remedial and 23 percent in advanced courses (NCES 1982, p. 78). By contrast, in 1972, according to Fetters, Brown, and Owings (1984, p. 13), only 6 percent of all seniors had taken remedial English courses in high school, and only 4 percent had taken remedial mathematics.

CHOICE OF CURRICULUM

The Score Decline Panel (College Board 1977, p. 41), writing in 1977, noted that "test takers enrolled in 'career' courses of study in high school...average much lower scores on the SAT than do those enrolled in 'academic' courses."

Since academically abler students gravitate toward academic courses, one would expect students in the academic curriculum to do better on the SAT if only as a consequence of self-selection. Alexander and Pallas (1984) have shown that there is more to the story. In a large scale study, they found that able students who did well on SCAT in grade 9 and who subsequently took a more academic program in high school made higher test scores in grade 12 than did a matched group of pupils who followed a less academic curriculum.

There is now evidence to show that in the 1970s there was a movement of high school students away from academic curriculums and toward vocational and general high school courses. Fetters, Brown, and Owings (1984) report that from 1972 to 1980 the percentage of seniors who considered themselves to be enrolled in an academic or college preparatory program declined from 46 to 38 percent. This shift may well reflect increasingly pragmatic or materialistic attitudes among students during the 1970s. The result, however, was probably a drop in the rigor of academic preparation which, in light of the study by Alexander and Pallas (1984), would directly affect SAT scores.

HOMEWORK

Although the research literature on homework is not unequivocal, most studies have shown that more hours of homework are associated with higher grades and test scores. Keith (1982) found, in fact, that with one to three hours of homework a week, the average student in the lowest 25 percent of the class in ability achieved grades commensurate with an average (middle 50 percent) student who did not do homework. Similarly, Keith's data show that a middle-ability student who did two to three hours more homework than a high-ability (top 25 percent) student received as good grades, on average, as the more able student.

In view of these findings, it is important to consider changes that have occurred in the amount of homework done by high school students over the years. By the end of the 1960s, the hearsay evidence about a declining trend in

homework was widespread. The Score Decline Panel reflected the same perception in 1977, noting "...the relatively clear evidence (though none of it has been quantified) of reduced assignments of homework" (College Board 1977, p. 30). Data from a major national 1984 study (Fetters, Brown, and Owings 1984) add some quantification: seniors' self-reports of the average amount of time per week spent on homework fell from 4.26 to 3.85 hours or by 25 minutes between 1972 and 1980. Whereas in 1972, 35.2 percent of seniors reported doing five hours or more a week; in 1980 only 24.5 percent reported doing as much.

The consensus of observer comment and the available data, then, suggest that at the high school senior level, homework began declining in the 1960s and continued to drop through the 1970s, as did the SAT scores. The downward trend in homework would appear to reflect a relaxation of demand, perhaps out of discouragement over the fact that homework is often not done even when assigned, perhaps out of concern that grading homework takes scarce teacher time. The extent to which the amount of homework done should be seen as something under the school's control is debatable: parents and students obviously play critical roles in any effective homework "policy." Yet this is a variable clearly within the overall responsibility of the school, which must set the standard of expectation. In a recent review of possible ways of increasing student attainment, Walberg (1984) found that the impact of homework on learning can be substantial in comparison with other variables. He points out that "homework that is graded or commented upon has three times the effect of SES. By comparison, homework that is merely assigned has an effect comparable to SES." Thus the decline in homework emerges as a matter of some importance.

GRADE INFLATION

The Panel on Score Decline had noted (College Board 1977, p. 29) that, in 1977, "An 'A' or 'B' means a good deal less than it used to." Evidently the trend toward easier grading standards did not end when the major compositional shift in the senior class had largely run its course in the early 1970s. A comparison of self-reported grades of high school seniors in 1972 and 1980 shows that 28.4 percent of the grades were A's and B's in 1972 but rose to 32.5 percent in 1980 (NCES 1982, p. 76). It is notable that throughout the 1970s, as in the 1960s, grades continued to rise just as inexorably as the test scores continued to fall. It seems reasonable to conclude that students were being given A's for work that would have drawn B's 10 years earlier. Students who had no reason to be aware of the change would have every right to accept their A's as *bona fide* evidence of top performance—and so would their parents. Relaxation of effort in class and of exhortation at home would be predictable consequences.

FEWER HIGH SCORES

On Further Examination took note of one phenomenon that seemed discordant with the compositional shift hypothesis, which neatly explained the appearance of more *low* scores on the basis of the retention of more low performers. The phenomenon that it failed to explain was the growing scarcity of *high* scores. The problem was described as follows in an appendix by Rex Jackson (1976a):

> One of the more mysterious aspects of the decline in verbal and mathematical scores by Scholastic Aptitude Test (SAT) candidates has been the decline in numbers of high scoring students. Had the number of high scoring students not declined, a theory attributing declines in the mean scores to an enlargement of opportunity for college attendance and consequent changes in the composition of the SAT candidate group would seem more plausible. Since the declines in mean SAT scores are not attributable solely to increased numbers of low scoring students, but also to smaller numbers of high scoring students, it appears that such a theory cannot fully explain the decline in scores (p. 1).

Jackson's analysis was not readily incorporated into the body of the Panel's report (College Board 1977), which notes:

> Some of the information initially submitted to the panel suggested an aberrational drop since 1970 in the number of students scoring 600 or above on either the Verbal or the Mathematical sections of the test or on both. That "high scorer" figure which stood at 189,300 in 1970 (among *all* SAT takers, as distinguished from a one-year cohort), had gone down by 1976 to 108,200; and this drop seemed to suggest elements of compositional or other related change.
>
> An inordinately extended analysis of the drop in the number of these "high scorers" indicates, however, that the decline probably results almost entirely from the reduction in the number of students taking the SAT and from the impact at the top of the same pervasive influences that have been affecting the scores of the test-takers as a whole (pp. 19-20).

The Panel, although a little troubled, did not pursue the matter further, but the decline in high scores has continued and remains a source of concern.

SCORES OF HIGH-RANKING AND LOW-RANKING STUDENTS

One might predict that if the instructional pace came to be set more by students toward the low end of the achievement level, there would be a tendency for students attaining a high rank in class to make lower scores. This should be compensated for in some degree by increases in the scores of students with a low class rank. This seems, in fact, to have happened, at least since 1976 when the College Board first began reporting the relevant data (College Board

1976–1980). Table 2 shows the figures in pivotal years for verbal and mathematical scores averaged together. Between 1976 and 1980, when SAT mean scores (verbal and mathematical combined) dropped seven points overall, the SAT mean for students who reported their class rank as in the top tenth dropped 13 points, or about twice the average for the total group. At the same time the students in the lowest fifth of the class improved their SAT scores by an average of six points.

By contrast, in the modestly up years of 1980 to 1984, the greatest gains were made by the students just below the top rank (second tenth and second fifth, or 60th to 90th percentile), who gained seven points while the top tenth gained four and the lowest fifth held even.

One can easily read too much into these data. The group from the lowest fifth of the class that takes the SAT, in particular, is a small and variable group comprising only about 1 percent of SAT takers. Still, the generalization would seem to be that the SAT scores of students at or near the top of the class dropped the most in the late 1970s while the scores of students toward the bottom were holding steady or gaining. In the 1980s, gains were spread broadly across class ranks, with some hint that the upper-middle class ranks were gaining the most.

These findings are in general agreement with the hypothesis that the instructional pace of the late 1970s was geared more nearly to the needs of the lower-achieving students. The trend since 1980 seems to have been more broadly beneficial—with some suggestion that it may not be helping either the highest or lowest students as much as those in the middle.

EARLIER GRADES

It has been noted earlier that the score decline of the 1960s and 1970s was not unique to the SAT, but characterized

Table 2. Mean SAT Scores by Class Rank (1976, 1980, 1984)

			Change		Change
Year	*1976*	*1980*	*'76 to '80*	*1984*	*'80 to '84*
Top Tenth	552	539	−13	543	+4
Second Tenth	478	470	−8	477	+7
Second Fifth	434	431	−3	438	+7
Third Fifth	386	386	0	392	+6
Fourth Fifth	361	360	−1	365	+5
Lowest Fifth	347	353	+6	353	0
Total	452	445	−7	449	+4

N.B. Total score means refer to all SAT takers. All other score means refer to those SAT takers who answered the Student Descriptive Questionnaire. The responding subgroup was approximately 77 percent of all test takers in 1976, 91 percent in 1980, and 90 percent in 1984. Class ranks are self-reported.

other data sets like ACT and NAEP scores gathered over time and *relating specifically to students in the last year or two of high school.* People who found "compositional change" a persuasive explanation for the first leg of the decline noted that the holding power of the schools had been strong at the lower grades for many years and so their student bodies had always been a more or less complete cross section of the age group. It was in grade 11 and especially in grade 12 that the student group had experienced selective dropout in earlier years. In the 1960s, it was the senior high school that came to retain a wider cross section of students and became notably more heterogeneous. This set of observations explains the first part of the decline nicely, especially the increase in low scores in the senior year but not in the early grades. The compositional-effect hypothesis for the first part of the decline seems to fit the facts relating to age and grade fairly well.

By contrast, it is hard to see why the "pervasive influences," independent of compositional effects, should lead to a selective continuation of the decline peculiar to the grades in which the compositional shift had occurred. This creates a problem for any "unrelated declines" explanation. According to the "reciprocal score declines" hypothesis, however, it was the new heterogeneity above grade 10 that led the schools to respond with a variety of pedagogical changes in grades 11 and 12 and those changes, in turn, fed the 1970s drop in performance in the grades affected. The earlier grades were less involved in either the compositional change or the pedagogic response to it and their scores reflected their stability.

COMPOSITIONAL CHANGE HAD A PEDAGOGICAL ECHO

The hypothesis I am advancing is that the profound change of the 1960s in the school retention rates for low-achieving students was the most powerful driving force behind the score declines of the 1970s and early 1980s. A brief and oversimplified account of what may have happened may serve as a summary of the "reciprocal declines" hypothesis.

In the 1960s a new wave of students whose academic performance was marginal "bought" the idea of staying in school to get a diploma. This was not a sudden change but rather a steadily growing acceptance of the society's conviction that a high school education was the minimum requirement for a good job. Employers both accepted and reinforced that proposition, as did the schools. Increasingly, students both accepted it and acted on it, with their parents' support.

In the 1970s the schools made their adjustment to the new facts of life. They now had a group of students staying past the legal school-leaving age whose preparation was weak by historical standards and by comparison with their classmates. The schools had to either flunk large numbers of these "new" senior high school students or inflate grades in

fairness to the others, and they chose the latter course. Everybody's grades moved up. Schools had to either assign material that was over the heads of a large proportion of the class or adopt easier textbooks: with the help of the publishers, they moved to easier textbooks. Schools also added electives that would accommodate students who found regular courses too difficult. They observed that homework was not done and they stopped requiring it.

The revolution of falling expectations was not lost on the students. What used to be B or B+ work was now rating an A or A+. Both the students and their parents were content with the visible evidence of accomplishment: the heat was off. Meanwhile, the textbooks that students were coping with in grade 12 had been reduced in difficulty by a grade or more. In grade 12 the students were meeting every demand being made of them—but the demand was at the previous 11th grade level. *Mirabile dictu:* their SAT scores looked remarkably like those of 11th grade students 10 to 20 years earlier.

This scenario suggests that while the 1960s leg of the decline was caused *directly* by compositional change, the 1970s leg reflected program changes that constituted the schools' developing response to that same change. Hence the "reciprocal score declines" hypothesis.

WHAT CAN THE DECLINE TEACH US?

The reason for revisiting the score decline issue lies in the potential importance of a change of such magnitude for educational theory, policy, and practice. Although SAT scores are not a good index of school performance, they are not irrelevant to the state of learning in this country. They are a symptom that something has disturbed and continues to affect an important segment of our learning system. And yet we seem in recent years to have come to a dead end in our exploration of the etiology of the malady, which is education's equivalent of the one-time mystery of "Legionnaires' Disease."

Various commissions have prescribed cures for the largely undiagnosed problem. It would seem important that we continue to pursue the search for a better understanding of what ails us as a guide to what to do about it. An important reason for examining the "pedagogical response" explanation for the second leg of the score decline is that it accords primacy to school factors—to things we can do something about if the hypothesis withstands scrutiny.

THE DECLINE IS PROBABLY OVER

Since 1980 we have seen the barest reversal of the SAT decline. Verbal score means are up by two points and mathematical score means by five. Only hindsight after a few more years will tell us for sure whether or not the long slide ended and reversed course in 1980. And yet, if the "reciprocal declines" hypothesis is correct, there are valid reasons to suppose that the trend is once again upward.

The "pedagogical response" to greater heterogeneity in senior high school has now had 15 or 20 years to effect the adjustments that it entailed. The consequences in student performance became alarmingly clear to the schools and the public nine or ten years ago, most vividly in the score decline data. Educational quality has had the national spotlight turned on it. The national mood turned strongly in support of a variety of changes, which included not only "back to basics" in the narrow sense, but more broadly a new emphasis on rigor.

The studies of high schools that have come along recently serve the important purpose of authenticating not only the concern but also the general directions in which trends are occurring. The changes seem likely to accelerate the pace of change. Concerted efforts are now being made to ensure the production and the selection of more challenging textbooks. There is more emphasis on the core curriculum and on "hard electives" such as Advanced Placement courses. No evidence has surfaced recently to suggest that grade inflation is increasing further and there are some indications from published College Board data that the trend has at least been arrested. Given these developments, backed by supportive public opinion, it seems reasonable to expect that the major trends leading to the decline have already been halted and probably reversed.

CONCERN ABOUT OVEREMPHASIS ON ACADEMIC PROFICIENCY

The concern that many people have is that we will buy better scores at the expense of the students for whom a rigorous academic emphasis is not the main aim in high school. The public expects the schools to provide not only academic learning for the college-bound but also the general education and career preparation that the other students (half of the class) want and need. Higher test scores can easily become an overriding goal that would deny the diversity of the teenage population and of the purposes of education in a democracy. This point needs to be made forcefully as a counter to an exclusive concern for academic ability in general and high-scoring students in particular.

Meanwhile, it is of interest that the students are quite clear on their own opinion as to the main change that is needed. In their comparison of student attitudes in 1972 and 1980, Fetters, Brown, and Owings (1984, p. 17) quote the evaluations of the seniors themselves as follows: "The largest change observed was an increase from 50 to 72 in the percent of students who felt that their school 'should have placed more emphasis on basic academic subjects (mathematics, science, English, etc.).'"

NEEDED: A SYSTEM TO SERVE DIVERSITY

A vigorous effort simply to reverse the changes in school practice of the past 20 years that may have contributed to the score decline might well succeed in getting us back to where we were educationally in the early 1960s. This hardly seems a stirring prospect in 1985. We need a new conception of how to reshape education to attain a high quality of student performance across the full spectrum of abilities and interests that has characterized the senior class since the mid-1960s. Since the provision of equal schooling for all subgroups of the population is still incomplete, we need to plan for serving a still greater diversity of students well in years to come.

This report is concerned principally with what led to the score declines of two decades. It would be out of place here to speculate on the outlines of what a new conception might entail. If it is to serve diversity, we have a host of little-used and promising approaches that can be tried for students showing a wide variety of talents and interests. Among these are breaking our insistence on the same number of years in high school for all students (teaching "to mastery" rather than to the calendar), utilizing acceleration and enrichment to better advantage, introducing hard electives to balance easy ones, instituting challenge exams or honors exams to complement minimum competency tests, and adapting technology to the educational needs of students in both advanced and remedial classes. All of these changes would work in the direction of more individualized instruction. The experience of the past 20 years suggests that the techniques we adopted may have accommodated the needs of less proficient students to the detriment of those of medium and high achievement. The need pointed out by the present investigation is to devise strategies that will allow us to attend to student needs in all their variety, taking explicit account of the new diversity of the senior high school class.

REFERENCES

Alexander, K.L., and A.M. Pallas, "Curriculum Reform and School Performance: An Evaluation of the 'New Basics,'" *American Journal of Education*, Vol. 92, No. 4, August 1984.

Chall, J.S., with S.S. Conrad and S.H. Harris, *An Analysis of Textbooks in Relation to Declining SAT Scores*, November 1976 (Revised March 1977). This report appears as an appendix to *On Further Examination*.

Chronicle of Higher Education, "Rises Reported in SAT Math, Verbal Scores," September 26, 1984.

College Board, *College Board News*. New York: College Entrance Examination Board, Spring 1984.

_____, *National College-Bound Seniors*. New York: College Entrance Examination Board, Annual Editions, 1976–1980.

_____, *On Further Examination*, Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline. New York: College Entrance Examination Board, 1977.

Fetters, W.B., G.H. Brown and J.A. Owings, *High School Seniors: A Comparative Study of the Classes of 1972 and 1980*. Washington, D.C.: National Center for Education Statistics, 1984.

Jackson, Rex, *An Examination of Declining Numbers of High-Scoring SAT Candidates*, October 1976(a). This report appears as an appendix to *On Further Examination*.

_____, *Mean Scores for PSAT and PSAT/NMSQT Junior Candidates*, 1976(b). This reports appears as an appendix to *On Further Examination*.

Keith, T.Z., "Time Spent on Homework and High School Grades: A Large-Sample Path Analysis," *Journal of Educational Psychology*, Vol. 74, 1982, pp. 248–253.

National Center for Education Statistics, *The Condition of Education*. Washington, D.C.: NCES, 1982, p. 78.

Turnbull, W.W., "Relevance in Testing," *Science*, Vol. 160, June 28, 1968.

Walberg, H.J., "Improving the Productivity of America's Schools," *Educational Leadership*, Vol. 41, No. 8, 1984, p. 24.